

AD-A103 427

NAVAL OCEAN SYSTEMS CENTER SAN DIEGO CA
THE MLE TRACKING ALGORITHM: A SUMMARY AND ERROR ANALYSIS. (U)
MAY 81 M J SHENSA
NOSC/TR-689

F/8 12/1

UNCLASSIFIED

NL

1 of 1
84-3037
SIC 3472

NOSC

END
DATE
FILED
4-8-81
DTIC

12

LEVEL II

RS

NOSC

NOSC TR 689

NOSC TR 689

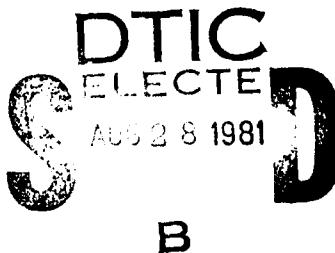
Technical Report 689

THE MLE TRACKING ALGORITHM: A SUMMARY AND ERROR ANALYSIS

M.J. Shensa

15 May 1981

Final Report: December 1980 – April 1981



Prepared for
Naval Sea Systems Command

Approved for public release; distribution unlimited

NAVAL OCEAN SYSTEMS CENTER
SAN DIEGO, CALIFORNIA 92152

50 8 28 00

DTIC FILE COPY



NAVAL OCEAN SYSTEMS CENTER, SAN DIEGO, CA 92152

AN ACTIVITY OF THE NAVAL MATERIAL COMMAND

SL GUILLE, CAPT, USN

Commander

HL BLOOD

Technical Director

ADMINISTRATIVE INFORMATION

Work for this report was performed from December 1980 to April 1981, under NAVSEA funds—Element 62633N; Project F33341; Task Area SF33341401; Work Unit 632-WC17.

Reviewed by
P. M. Reeves, Head
Electronics Division

Under authority of
R. H. Hearn, Head
Fleet Engineering Department

ACKNOWLEDGMENT

The author wishes to thank J. Pack for his help in implementing the software used in this analysis.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE			READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NOSC Technical Report 689 (TR 689)	2. GOVT ACCESSION NO. AD-1103427	3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) THE MLE TRACKING ALGORITHM: A SUMMARY AND ERROR ANALYSIS.	5. TYPE OF REPORT & PERIOD COVERED Final: December 1980 - April 1981		
7. AUTHOR(s) M. J. Shensa	8. CONTRACT OR GRANT NUMBER(s)		
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Ocean Systems Center San Diego, CA 92152	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62633N/F33341/SF33341401, 4 632-WC17		
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Sea Systems Command Washington, D.C.	12. REPORT DATE 15 May 1981		
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	13. NUMBER OF PAGES 26		
16. DISTRIBUTION STATEMENT (of this Report)	15. SECURITY CLASS. (of this report) Unclassified		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)	15a. DECLASSIFICATION DOWNGRADING SCHEDULE		
18. SUPPLEMENTARY NOTES			
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) least squares tracking algorithms maximum likelihood estimator Kalman filters confidence ellipsoids			
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This report discusses, in general, least-squares tracking algorithms and in particular the maximum likelihood estimator (MLE). Its primary concern is a detailed error analysis both from a deterministic point of view (sensitivity study) and a stochastic viewpoint (state variable covariance matrix). The latter results are shown to be equivalent to a Cramer-Rao bound, and their relationship to Kalman filters is cited. Various subtleties of interpretation are discussed including several theorems on confidence ellipsoids. Examples generated by computer software based on the theory are also presented.			

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 68 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

S-N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

SUMMARY

This report discusses, in general, least-squares tracking algorithms and in particular the maximum likelihood estimator (MLE). Its primary concern is a detailed error analysis both from a deterministic point of view (sensitivity study) and a stochastic viewpoint (state variable covariance matrix). The latter results are shown to be equivalent to a Cramer-Rao bound, and their relationship to Kalman filters is cited. Various subtleties of interpretation are discussed including several theorems on confidence ellipsoids. Examples generated by computer software based on the theory are also presented.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distributive / _____	
Available to _____	
Approved by _____	
Dist	1. Civil
A	

CONTENTS

1. INTRODUCTION . . . page 5

2. SENSITIVITY AND ERROR ANALYSIS . . . 7

 The Role of the Base Frequency . . . 9

 Remarks . . . 10

 Cramer-Rao Bound . . . 11

APPENDIX A: COMPUTATION OF GRADIENT AND "JACOBIAN" . . . 15

APPENDIX B: STATE VARIABLE COVARIANCE AND ERROR ELLIPSES . . . 19

APPENDIX C: SOME THEOREMS ON POSITIVE DEFINITE MATRICES . . . 24

REFERENCES . . . 26

1. INTRODUCTION

This report is concerned with least-squares tracking algorithms predicted on minimizing frequency and bearing errors. The state variables which we wish to determine may be considered to be the initial target position (x-coordinate and y-coordinate) and target velocity for each segment of the target track (2(number of maneuvers + 1) variables). The bearing angles θ^t and frequencies f^t will be a function of the unknown state variables $\bar{q} = (q_1, q_2, \dots)$ as well as of parameters associated with the receiving platforms (position, velocity, etc.) which we denote by the vector \bar{z} . The superscript "t" is intended to suggest time, but can actually index any general set of measurements not necessarily in chronological order and possibly received by several platforms.

In terms of the above definitions, the problem may be expressed as:

Determine \bar{q} which minimizes

$$G(\bar{q}) = \sum_{t=1}^T a_t (\theta^t(\bar{q}, \bar{z}) - \hat{\theta}^t)^2 + \sum_{t=1}^T b_t \left(\frac{f^t(\bar{q}, \bar{z})}{f_0} - \frac{\hat{f}^t}{f_0} \right)^2 \quad (1)$$

where $\hat{\theta}^t$ and \hat{f}^t are measured quantities and a_t and b_t are suitable weights (set equal to zero in the absence of a measurement). Note that the source frequency f_0 can either be considered a parameter (such as \bar{z}) or one of the state variables \bar{q} . A necessary condition for $\hat{\bar{q}}$ to yield a minimum is that it satisfy the set of nonlinear equations

$$\nabla_{\bar{q}} G(\bar{q}) \Big|_{\bar{q}} = 0 \quad (2)$$

obtained by setting the gradient of (1) equal to zero.

An obvious choice for the weights is to make them proportional to the accuracy of the corresponding measurements. If we take $a_t = (\sigma_{\hat{\theta}^t})^{-1}$ and $b_t = (\sigma_{\hat{f}^t})^{-1}$ where σ^2 denotes variance we obtain the so-called maximum likelihood estimator (MLE) [1,2]. This nomenclature reflects the fact that if $\hat{\theta}^t$ and \hat{f}^t are independent Gaussian random variables, then the minimum $\hat{\bar{q}}$ is the maximum likelihood estimate of \bar{q} . Bowing to current usage [2], we shall continue to use the term MLE although it is a misnomer when a_t and b_t are not the true variances.

There are many standard techniques for solving (2). We merely note here that they typically involve the computations of the gradient and/or the Jacobian of $G([2], [3], [4])$. These computations are discussed in Appendix A. Our primary concern in this report is

1. Jazwinski, A., Stochastic Processes and Filtering Theory, Academic Press, N.Y., 1970.
2. Adams, W.B., "A Non-Linear Maximum *a posteriori* Estimator for Passive Ranging Using Bearing and Frequency Measurements on a Stable Line," General Electric Co., Syracuse, N.Y., Sep 1975.
3. Powell, M., "A Survey of Numerical Methods for Unconstrained Optimization," SIAM Review, Vol 12, pp 79-97, Jan 1970.
4. Bard, Y., "Comparison of Gradient Methods for the Solution of Nonlinear Parameter Estimation Problems," SIAM J. Num. Analysis, Vol 7, pp 157-185, March 1970.

the accuracy of the solution \hat{q} as a function of the statistics of $\hat{\theta}$, \hat{f} , and \bar{z} . This is treated in the next section. The analysis permits the direct calculation of the error statistics (to a reasonable approximation), a capability which is important in practical implementations and, as well, alleviates the need for Monte Carlo techniques in evaluations of the algorithm.

We shall also see that for Gaussian statistics, the estimated MLE error is essentially the Cramer-Rao bound [5] for the solution track. If the errors are not too large this estimate a good approximation of the actual error statistics; i.e., the MLE achieves minimum variance compared to all unbiased estimators (with white Gaussian noise). A similar result has been shown for the Kalman filter* linearized about the true solution [6]. Thus, if one assumes the bias is negligible (as it appears in practice), the local accuracy of the MLE and extended Kalman filter should be approximately the same, and the error analysis developed here should serve as a lower bound for this entire class of algorithms.

Furthermore, the computations in the following section include a sensitivity analysis with respect to most of the significant measured parameters such as platform position, speed, etc. (designated \bar{z} in equation (1)). The validity of these results extends to any algorithm designed to minimize the MLE cost function, and, thus, should also be indicative for the various nonlinear Kalman filters.

In brief, we shall first derive general expressions for the error by linearizing and taking expectations. These are then specialized to cost functions of the type (1). Specifically, we distinguish between the roles of the random variables \bar{z} and those of $\hat{\theta}$ and \hat{f} . Finally, we show that this procedure leads to an estimate of the variances of the state variables equal to the Cramer-Rao bound. Appendices A and B contain explicit computations of equation (2), the Jacobian of that system of equations evaluated at the solution point, and the covariance matrix for the state variables. Appendices B and C also present a fairly detailed analysis of the behavior of the corresponding confidence ellipsoids. Various remarks interpreting the theory and pointing out its (sometimes) subtle consequences are scattered throughout the text and are considered to be one of the main contributions of this report. In particular, the examples found in Appendix B should not be overlooked.

*For linear systems, the Kalman filter (with unknown initial state) coincides with the MLE [1]; however, nonlinear versions such as the Iterated Extended Kalman Filter (IEKF) are, in a sense, only approximations to the MLE. The Kalman filter, being a recursive algorithm, generates new states utilizing only the previous state and the most recent observation. For linear estimators this is sufficient since the "linear least squares" or "minimum variance" correction is a linear combination of the observation and old state (orthogonal to the old state). For nonlinear estimators, the new estimate is not necessarily orthogonal to the old since the new estimation space is entirely different: it involves nonlinear combinations with past data and is, in general, infinite dimensional. In many situations the approximation involved in the IEKF are good; however, they can diverge, and the IEKF covariance estimate is often overly optimistic [1, 7].

1. Jazwinski, A., Stochastic Processes and Filtering Theory, Academic Press, N.Y., 1970.
5. Van Trees, H.L., Detection, Estimation, and Modulation Theory, Part I, Wiley, New York, N.Y., 1968.
6. Taylor, J., "The Cramer-Rao Estimation Error Lower Bound Computation for Deterministic Nonlinear Systems," Proceedings of 1978 IEEE Conference on Decision and Control, pp 1178-1181.
7. Chou, S.I., "Some Drawbacks of Extended Kalman Filters in ASW Passive Angle Tracking," ONR Symposium on Advances in Acoustic Passive Tracking, NPS-62 TS 77771, Naval Post Graduate School, Monterey, CA, May 1977.

2. SENSITIVITY AND ERROR ANALYSES

Let \bar{q} be a vector representing the set of state variables (target position, velocity, etc.) and let w be a (scalar) measured parameter such as receiver position or target bearing measurement. Assume that we are solving a system of equations in \bar{q} of the form

$$F_k(\bar{q}, w) = 0 \quad k=1, \dots, N \quad . \quad (3)$$

The sensitivity to the measurement w may be computed by varying w from its true value \hat{w} and linearizing about the point (\bar{q}, \hat{w}) satisfying $F_k(\bar{q}, \hat{w}) = 0$. That is,

$$\sum_m \frac{\partial F_k}{\partial \bar{q}_m} (\Delta \bar{q})_m + \frac{\partial F_k}{\partial \hat{w}} \Delta w \sim 0 \quad (4)$$

or

$$\Delta \bar{q} \sim -J^{-1} \left(\frac{\partial F}{\partial \hat{w}} \right) \Delta w \quad , \quad (5)$$

where

$$\Delta \bar{q} = \bar{q} - \hat{q} \quad (6a)$$

$$\Delta w = w - \hat{w} \quad (6b)$$

$$J_{km} = \frac{\partial F_k}{\partial \bar{q}_m} \Bigg|_{(\hat{q}, \hat{w})} \quad (6c)$$

and F is an N -dimensional vector valued function, with components F_k .

More typically we are concerned with functions which are a sum of terms of the form (3)

$$\sum_t F_k^t (\bar{q}, w^t) = 0 \quad k=1, \dots, N \quad . \quad (7)$$

Let us assume that $\Delta w^t = w^t - \hat{w}^t$ for $t=1, \dots, T$, constitutes a set of zero-mean independent identically distributed random variables such as, for example, the set of bearing measurements at time t . Then we may write $E(\Delta w^t \Delta w^{t'}) = 0$ and $\sigma_w^2 = E(\Delta w^t)^2$. Taking the expectation of the product of components of (5), we obtain

$$E(\Delta \bar{q}_k \Delta \bar{q}_m) \sim \sigma_w^2 \sum_t \left(J^{-1} \frac{\partial F}{\partial w^t} \right)_k \left(J^{-1} \frac{\partial F}{\partial \hat{w}^t} \right)_m \quad (8a)$$

or, for $m = k$,

$$E(\Delta \bar{q}_k)^2 \sim \sigma_w^2 \sum_t \left(J^{-1} \frac{\partial F}{\partial \hat{w}^t} \right)_k^2 \quad (8b)$$

Let us now assume that equation (7) is the result of a nonlinear least-squares fit (cf., eq. (1)); i.e., arises from the minimization problem

$$\min_{\bar{q}} \frac{1}{2} \sum_t \left(c^t(\bar{q}, z^t) - \tilde{c}^t \right)^2 \quad (9)$$

where for simplicity of notation we take z^t to be a scalar. For the general case see Appendix B and equation (B-8). Both z^t and \tilde{c}^t represent measured parameters (for example, receiver speed and target bearing normalized by the least square weights respectively) and thus are examples of the variable w^t in (7). However, their roles are typically perceived as being physically different, and they are distinguished mathematically by the different functional dependencies of expression (9). More intuitively, we might say that the \tilde{c}^t are those parameters which have been chosen to "measure" the quality of fit of the least squares track.

The solution to (9) satisfies the set of equations

$$\sum_t (c^t - \tilde{c}^t) \frac{\partial c^t}{\partial \bar{q}_k} = 0 \quad k=1, \dots, N \quad (10)$$

Noting that at the point of linearization (\bar{q}, \bar{z}^t) and for the true measured value \hat{c}^t , the expression $c^t - \tilde{c}^t$ vanishes, we find the Jacobian of system (10) to be given by

$$J_{km}(\hat{q}, \hat{z}, \hat{c}) = \sum_t \frac{\partial c^t}{\partial \hat{q}_k} \frac{\partial c^t}{\partial \hat{q}_m} \quad (11)$$

Also,

$$\left(\frac{\partial F^t}{\partial \hat{z}^t} \right)_m = \frac{\partial c^t}{\partial \hat{z}^t} \frac{\partial c^t}{\partial \hat{q}_m} \quad (12a)$$

$$\left(\frac{\partial F^t}{\partial \tilde{c}^t} \right)_m = - \frac{\partial c^t}{\partial \tilde{q}_m} \quad . \quad (12b)$$

THE ROLE OF THE BASE FREQUENCY

In Doppler tracking the base frequency f_0 sometimes plays the role of a measured parameter w and sometimes that of an unknown state variable \bar{q}_m . In the former case, following equation (1), we may write $\tilde{c}^t = \sqrt{b_t} (\tilde{f}^t/f_0)$. The \tilde{c}^t are no longer independent, and we find (with $\tilde{f}^t/f_0 \sim 1.0$) that

$$\begin{aligned} E(\Delta \bar{q}_k)^2 &= \frac{\sigma_{f_0}^2}{f_0^2} \left(\sum_t J^{-1} \frac{\partial c^t}{\partial \tilde{q}} \right)_k^2 \\ &= \frac{\sigma_{f_0}^2}{f_0^2} \left(\sum_m J_{km}^{-1} \sum_t \frac{\partial c^t}{\partial \tilde{q}_m} \right)^2 \quad . \end{aligned} \quad (13)$$

In the second case, although the cost function does not have the form (9), if one takes $c^t = \left(\frac{f^t}{f_0} - 1 \right) \sqrt{b_t}$, the equation corresponding to $\bar{q}_m = f_0$ becomes

$$\sum_t \left[c^t - \left(\frac{\tilde{f}^t}{f_0} - 1 \right) \sqrt{b_t} \right] \left(\frac{\tilde{f}^t}{f_0^2} \right) \sqrt{b_t} = 0 \quad (14)$$

and relations (11) and (12) read

$$\begin{aligned} J_{k, f_0} &= \sum_t \frac{\hat{f}^t}{f_0^2} \frac{\partial c^t}{\partial \tilde{q}_k} \sqrt{b_t} \\ J_{f_0, f_0} &= \sum_t \left(\frac{\hat{f}^t}{f_0^2} \right)^2 b_t \end{aligned} \quad (15)$$

$$\begin{aligned} \left(\frac{\partial F^t}{\partial \hat{z}^t} \right)_{f_0} &= \frac{\partial c^t}{\partial \hat{z}^t} \quad \frac{\hat{f}^t}{\hat{f}_0^2} \sqrt{b_t} \\ \left(\frac{\partial F^t}{\partial \tilde{f}^t} \right)_{f_0} &= - \frac{\hat{f}^t}{\hat{f}_0^2} \quad b_t \end{aligned} \quad (16)$$

We note that the last equation is equivalent to $\left(\frac{\partial F^t}{\partial \tilde{c}^t} \right)_{f_0} = - \frac{\hat{f}^t}{\hat{f}_0^2}$. All derivatives were evaluated, of course, at the point representing the true parameter values.

Finally, in either case, if we are interested in the errors due to the measurements \tilde{f}^t , the appropriate relationship is (8) with σ_w^2 replaced by $\frac{\sigma_f^2}{\hat{f}_0^2} \triangleq \frac{\sigma_{\tilde{f}^t}^2}{\hat{f}_0^2}$ and $\frac{\partial F^t}{\partial \hat{w}^t}$ given by (12b) or (16).

REMARKS

1. Note that the entire analysis rests on the approximation (4). Although experience would lead one to believe that (4) is a valid step (for sufficiently small σ_x) in the derivation of (8), one must still retain a degree of skepticism. For example, taking the expectation of (4) yields $E(\Delta \bar{q}_m) = 0$; i.e., seems to imply that the least-squares estimates are unbiased. This is only true to the first order; and, as is well known, for many cost functions the bias can be significant (cf. [8] and the Churn algorithm of [7]).

2. Under the assumption that the various types of measurement errors are independent of each other, we may calculate $E(\Delta \bar{q})^2$ due to all these sources simply by summing the individual contributions (see Appendix B). The RMS error is then found by taking the square root.

3. Relation (11) which expresses the Jacobian at the "true solution" point holds at the solution for any minimization problem of the form $\min \Sigma(G^t)^2$ provided the minimum is zero.

4. The above results easily generalize to cost functions which are the sum of functions of the form (9); e.g.,

$$\min_{\bar{q}} \frac{1}{2} \sum_t [(c^t - \tilde{c}^t)^2 + (d^t - \tilde{d}^t)^2]$$

7. Chou, S.I., "Some Drawbacks of Extended Kalman Filters in ASW Passive Angle Tracking," ONR Symposium on Advances in Acoustic Passive Tracking, NPS-62 TS 77771, Naval Post Graduate School, Monterey, CA, May 1977.
 8. Eykhoff, P., System Identification, Chapter 6, John Wiley & Sons, N.Y., 1974.

where c and d represent bearing and Doppler measurements respectively. Also, nowhere has our analysis restricted us to measurements made from a single receiving platform. However, if the measurements have different variances, additional terms of the above form might be desirable.

CRAMER-RAO BOUND

Let us assume that the least squares weights have been chosen equal to the inverse variances of the measured quantities (true MLE) so that the variances of the measured parameters $(\sigma_{\tilde{c}^t})^2$ of expression (9) are equal to one.* If these variables are independent Gaussian variables, the likelihood function, up to a constant, is given by

$$L(q) = -\frac{1}{2} \sum_t (c^t - \tilde{c}^t)^2 \quad (17)$$

The Fisher information matrix is [5]

$$\begin{aligned} J_{km}^F &= E \left(\frac{\partial L}{\partial \tilde{q}_k} \right) \left(\frac{\partial L}{\partial \tilde{q}_m} \right) \Big|_{\hat{q}} \\ &= E \left[\sum_t \frac{\partial^2 c^t}{\partial \tilde{q}_k \partial \tilde{q}_m} (\hat{c}^t - \tilde{c}^t) \right] + \sum_t \frac{\partial c^t}{\partial \tilde{q}_k} \frac{\partial c^t}{\partial \tilde{q}_m} \\ &= \sum_t \frac{\partial c^t}{\partial \tilde{q}_k} \frac{\partial c^t}{\partial \tilde{q}_m} \\ &= J_{km} \end{aligned} \quad (18)$$

Thus, the Cramer-Rao bound, which is true for any unbiased estimate of q_k [5], may be expressed as

$$E(\Delta \tilde{q}_k)^2 \geq (J^{-1})_{kk} \quad (19)$$

On the other hand, our sensitivity analysis led to equation (8b), which with $\sigma_w^2 = \sigma_{\tilde{c}^t}^2 = 1$, gives

*As stated previously, these parameters are also assumed independent. The more general case may be treated by using the inverse covariance matrix in lieu of the weights.

S. Van Trees, H.L., Detection, Estimation, and Modulation Theory, Part I, Wiley, New York, N.Y., 1968.

$$\begin{aligned}
E(\Delta\bar{q}_k)^2 &\sim \sum_t \left(J^{-1} \frac{\partial F^t}{\partial \tilde{c}^t} \right)_k^2 \\
&= \sum_t \left(\sum_m J_{km}^{-1} \frac{\partial c^t}{\partial \bar{q}_m} \right)^2 \quad (\text{from 12b}) \\
&= \sum_t \sum_m \sum_{m'} J_{km}^{-1} J_{km'}^{-1} \frac{\partial c^t}{\partial \bar{q}_m} \frac{\partial c^t}{\partial \bar{q}_{m'}} \\
&= \sum_m \sum_{m'} J_{km}^{-1} J_{km'}^{-1} J_{mm'} \\
&= J_{kk}^{-1} \quad . \quad (20)
\end{aligned}$$

Thus, in this case, our estimated error of the MLE algorithm equals the Cramer-Rao bound. The same computation applied to (8a) also yields

$$Q = J^{-1} \quad (21)$$

where Q is the covariance matrix of $\Delta\bar{q}$.

Note that each new data point results in the addition of a matrix of the form $C_{km} = \frac{\partial c^t}{\partial \bar{q}_k} \frac{\partial c^t}{\partial \bar{q}_m}$. Since C is positive semi-definite, J is a non-decreasing matrix function ($J \geq J'$ if $J-J'$ is positive semi-definite) of the data. We conclude from Appendix C that the corresponding concentration ellipsoids $(\Delta\bar{q})^\dagger Q(\Delta\bar{q}) \leq 1$ (Appendix B and [5]) form a non-increasing nested sequence. In practice C is typically positive and the sequence is strictly decreasing. The reader should be warned, however, that although this represents an improved estimate of the state variables, it does not necessarily imply increased accuracy in the current target position. The position state variables are the coordinates of the target at a fixed point in time (for example $t=0$), and the errors must be propagated along the target track as detailed in Appendix B.

A second comment is that relations (20) and (21) assume that the *a priori* estimates of the measurement variances are correct; i.e., $a_t = 1/\sigma_\theta^2$ and $b_t = f_0^2/\sigma_f^2$ in equation (1). If this is not true,

$$E(\Delta\bar{q}_k)^2 = \sum_t \sum_m \sum_{m'} J_{km}^{-1} J_{km'}^{-1} (\sigma_{\tilde{c}^t})^2 \frac{\partial c^t}{\partial \bar{q}_m} \frac{\partial c^t}{\partial \bar{q}_{m'}} \quad (22)$$

and further reduction is not generally possible. These comments are illustrated in figures 1a and 1b. Inclusion of the errors due to the parameters z^t of (9) could have similar effects; in particular, $Q \neq J^{-1}$.

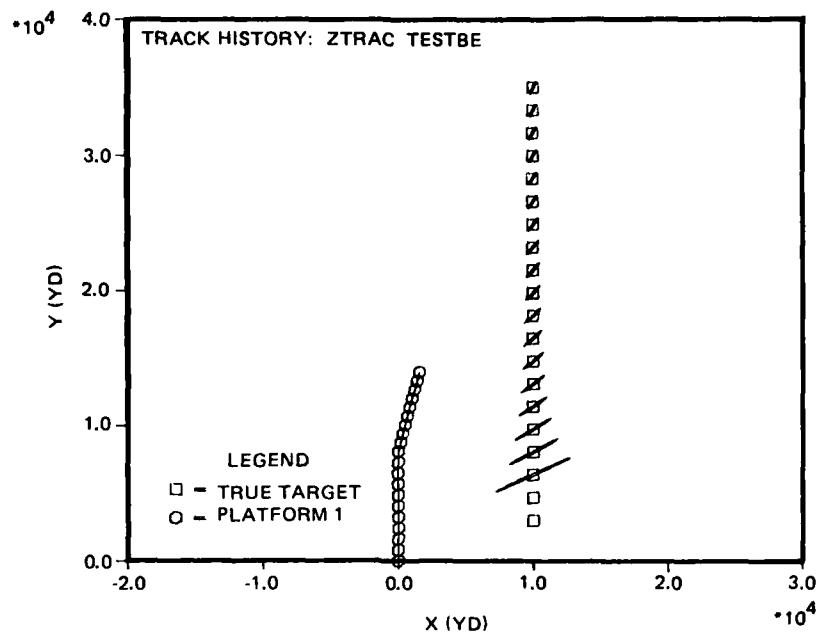


Figure 1a. Example illustrating diminishing error ellipses. Platform track course is 20 minutes at 0° and 12 knots; and 20 minutes at 15° , 10 knots. Target is 0° , 25 knots. Bearing standard deviation is $\sigma_\theta=0.4^\circ$ and frequency standard deviation is $\sigma_f=0.01$ Hz with a base frequency of 100 Hz. The MLE weights were taken to be σ_θ^{-2} and $(\sigma_f/f_0)^{-2}$, respectively. (The first two points have infinite error ellipses.)

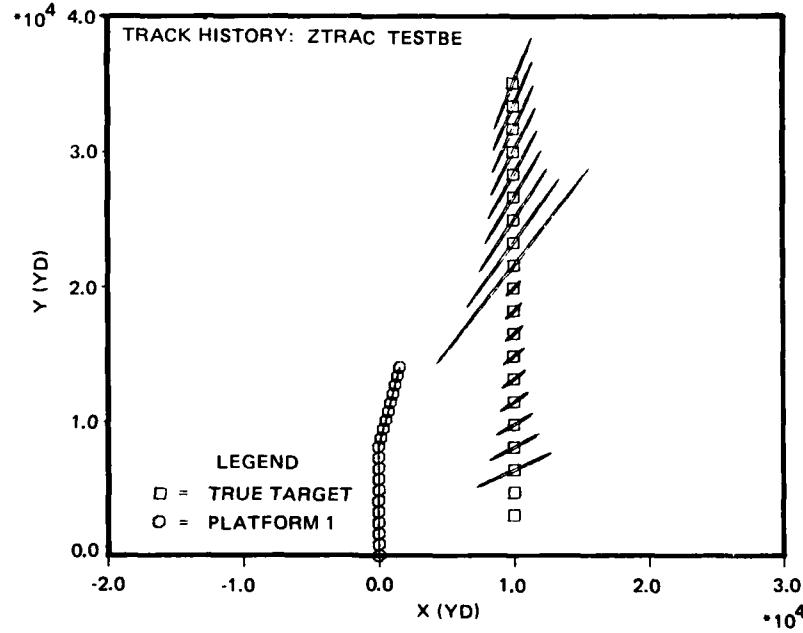


Figure 1b. The same as Figure 1a, only the MLE weights were taken to be σ_θ^{-2} and $(0.5/f_0)^2$ which corresponds to a σ_f of 0.5 Hz. Note that only the ratio of the two weights is relevant.

APPENDIX A COMPUTATION OF GRADIENT AND "JACOBIAN"

Let $\bar{p}(t) = (p_x(t), p_y(t))$ and $\bar{v}(t) = (v_x(t), v_y(t))$ represent the position and velocity of one of the receiving platforms at time t . Similarly let $\bar{p}'(t)$ and $\bar{v}'(t)$ be those of the source (target). Their relative position and velocity are given by:

$$\begin{aligned}\Delta\bar{p} &= \bar{p}' - \bar{p} \\ \Delta\bar{v} &= \bar{v}' - \bar{v}\end{aligned}\quad (A-1)$$

so that, if we define the positive y axis as north, the relative bearing of the target is given by (cf. Figure A.1)

$$\theta = \tan^{-1} \frac{(\Delta\bar{p})_x}{(\Delta\bar{p})_y} \quad (A-2)$$

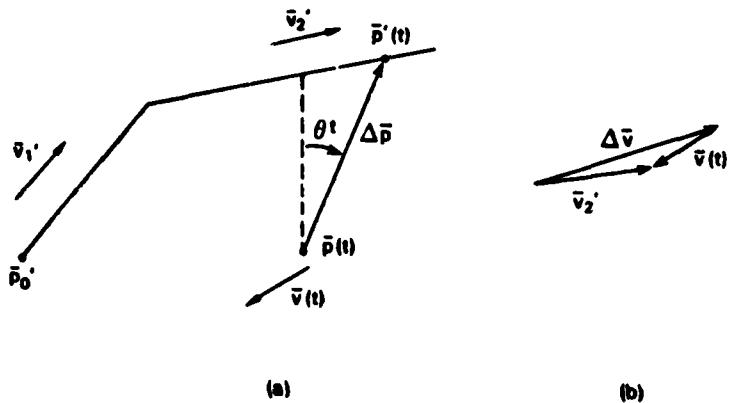


Figure A.1. Example of the geometry for a two segment target.

It is convenient (and provides improved roundoff characteristics) to deal with the normalized Doppler shift

$$d = \frac{f - f_0}{f_0} \quad , \quad (A-3)$$

where f_0 and f are the source and received frequencies respectively. If the velocities are small with respect to that of sound (denoted c), we may consider only the relative motion in computing d . This results in the following expression:

$$d = - \frac{\Delta \bar{v} \cdot \Delta \bar{p}}{c \parallel \Delta \bar{p} \parallel} \quad (A-4)$$

where "·" indicates scalar product.

We now assume that the target follows a piecewise linear track with initial position \bar{p}'_0 and velocity \bar{v}'_i over segment i from time t_i to t_{i+1} . Then

$$\bar{p}'(t) = \bar{p}'_0 + \bar{v}'_1(t_2 - t_1) + \cdots + \bar{v}'_i(t - t_i) \quad (A-5)$$

for $t_i \leq t < t_{i+1}$

Our state variables are taken to be $(\bar{p}'_0)_x, (\bar{p}'_0)_y, (\bar{v}'_i)_x, (\bar{v}'_i)_y$ where $i=1$ to the number of segments. The base frequency f_0 may also be a state variable.

The cost function to be minimized takes the form

$$G = \frac{1}{\sigma_\theta^2} B + \frac{1}{\sigma_d^2} D \quad (A-6)$$

where

$$B = \frac{1}{2} \sum_t (\theta^t - \tilde{\theta}^t)^2_{2\pi} \quad (A-7)$$

$$D = \frac{1}{2} \sum_t (d^t - \tilde{d}^t)^2 \quad (A-8)$$

and σ_θ^2 and σ_d^2 are the *a priori* variances of θ and d (for notational simplicity assumed constant functions of t). The symbol $\tilde{\theta}$ represents measured bearing and $\tilde{d} = \frac{\tilde{f} - f_0}{f_0} = \frac{\tilde{f}}{f_0} - 1$

where \tilde{f} is the measured frequency. The base frequency f_0 is either determined *a priori* from the data or is a state variable to be included in the minimization. The 2π in equation (A-7) indicates that the bearing difference is taken modulo 2π so as to yield the minimum difference, i.e., a value between $-\pi$ and π .

The partial derivatives of B and D with respect to a state variable q are

$$(B)_q = \sum_t (\theta^t - \tilde{\theta}^t)_{2\pi} (\theta^t)_q$$

$$(D)_q = \sum_t (d^t - \tilde{d}^t) (d^t - \tilde{d}^t)_q \quad (A-9)$$

with the subscript q indicating partial differentiation. These expressions may be evaluated using the chain rule. Thus, dropping the superscript t for simplicity, we have

$$(\theta)_q = \begin{cases} (\theta)_{\bar{p}'} \cdot (\bar{p}')_q & q \neq f_0 \\ 0 & q = f_0 \end{cases} \quad (A-10)$$

$$(d - \tilde{d})_q = \begin{cases} (d)_{\bar{p}'} \cdot (\bar{p}')_q + (d)_{\Delta \bar{v}} \cdot (\Delta \bar{v})_q & q \neq f_0 \\ \frac{\tilde{f}}{f_0^2} & q = f_0 \end{cases} \quad (A-11)$$

The “•” stands for scalar product; a vector subscript indicates the gradient; and the derivative of a vector is the vector comprised of the derivatives of its components.

The gradient vector $(\theta)_{\bar{p}'}$ is easily found from (A-2).

$$\begin{aligned} (\theta)_{p'_x} &= \frac{(\Delta \bar{p})_y}{\|\Delta \bar{p}\|^2} \\ (\theta)_{p'_y} &= \frac{-(\Delta \bar{p})_x}{\|\Delta \bar{p}\|^2} \end{aligned} \quad (A-12)$$

Likewise, (A-4) implies

$$\begin{aligned} (d)_{\bar{p}'_x} &= \frac{-1}{\|\Delta \bar{p}\|} \left[d \frac{(\Delta \bar{p})_x}{\|\Delta \bar{p}\|} + \frac{(\Delta \bar{v})_x}{c} \right] \\ (d)_{\bar{v}'_x} &= \frac{-(\Delta \bar{p})_x}{c \|\Delta \bar{p}\|} \end{aligned} \quad (A-13)$$

and similarly for the y components.

The other partials follow immediately from (A-5). Some of these values depend on the time t. The only “tricky” calculation is

$$\frac{\partial \bar{p}'_x}{\partial (\bar{v}'_i)_x} = \begin{cases} 0 & t \leq t_i \\ \frac{1}{\sigma_{\theta}^2} (\theta^t)_k (\theta^t)_m & t_i < t \leq t_{i+1} \\ \frac{1}{\sigma_d^2} (d^t)_k (d^t)_m & t_{i+1} \leq t \end{cases} \quad (A-14)$$

We also note that $(\Delta \bar{v})_q$ is zero unless $q = (\bar{v}'_i)_x$ or $(\bar{v}'_i)_y$ where $t_i \leq t < t_{i+1}$, in which case it is the vector (1,0) or (0,1) respectively.

The computation of J of equation (11) is given by

$$J_{km} = \sum_t \left[\frac{1}{\sigma_{\theta}^2} (\theta^t)_k (\theta^t)_m + \frac{1}{\sigma_d^2} (d^t)_k (d^t)_m \right] \quad (A-15)$$

Note that this is the true Jacobian of the system of equations $\nabla G = 0$ only at a point for which $\theta^t = \hat{\theta}^t$ and $d^t = \hat{d}^t$: i.e., only at a true solution with perfect measurements. Since, in practical situations, this is never the case, J only represents a (possibly poor) approximation to the Jacobian. Nevertheless, it is quite suitable for use in the steepest descent type algorithms employed in solving (2) [2].

Finally, we remark that if the various measurements have different accuracies, i.e., if $\text{var}(\hat{\theta}^t)$ is a function of t , then σ_{θ}^2 and σ_d^2 must be replaced by $\sigma_{\theta}^2(t)$ and $\sigma_d^2(t)$, which reside inside the summation over t .

2. Adams, W.B., "A Non-Linear Maximum *a posteriori* Estimator for Passive Ranging Using Bearing and Frequency Measurements on a Stable Line," General Electric Co., Syracuse, N.Y., Sep 1975.

APPENDIX B STATE VARIABLE COVARIANCE AND ERROR ELLIPSES

The equations developed in Section 2 suffice to compute the covariance matrix of the state variables \bar{p}'_0 and \bar{v}'_i ; however, we are typically interested in the accuracy of the estimated current target position $\bar{p}'(t)$ (see equation (A-5)). The covariance of $\bar{p}'(t)$ follows directly from (A-5)

$$E(\Delta\bar{p}_r(t) \Delta\bar{p}_q(t)) = E\left[\left(\Delta\bar{p}_{0r} + \sum_i \Delta\bar{v}_{ir} \Delta t_i\right) \left(\Delta\bar{p}_{0q} + \sum_i \Delta\bar{v}_{iq} \Delta t_i\right)\right]. \quad (B-1)$$

where r and q stand for either x or y , and $\Delta t_i = t - t_i$ or $t_{i+1} - t_i$. Let the covariance matrix of the state variables be given by

$$\begin{aligned} Q_{rq}^{pp} &= E(\bar{p}_{0r} \bar{p}_{0q}) \\ Q_{rq}^{pv_i} &= E(\bar{p}_{0r} \bar{v}_{iq}) \\ Q_{rq}^{v_i v_j} &= E(\bar{v}_{ir} \bar{v}_{jq}) \end{aligned} \quad (B-2)$$

Then (B-1) becomes

$$E(\Delta\bar{p}_r(t) \Delta\bar{p}_q(t)) = Q_{rq}^{pp} + \sum_i \Delta t_i \left[Q_{rq}^{pv_i} + Q_{qr}^{pv_i} + \sum_j \Delta t_j Q_{rq}^{v_i v_j} \right]. \quad (B-3)$$

Note that the matrices Q may be computed from (8a), (see also equation (D-8)).

Let us label the three quantities evaluated in (B-3) by

$$\begin{aligned} \sigma_{xx} &= E(\Delta\bar{p}_x \Delta\bar{p}_x) \\ \sigma_{yy} &= E(\Delta\bar{p}_y \Delta\bar{p}_y) \\ \sigma_{xy} &= E(\Delta\bar{p}_x \Delta\bar{p}_y) \end{aligned} \quad (B-4)$$

The equation of the ellipse for the position variables is then [10]

$$\frac{\sigma_{yy}}{D} (\Delta x)^2 - \frac{2\sigma_{xy}}{D} (\Delta x) (\Delta y) + \frac{\sigma_{xx}}{D} (\Delta y)^2 = c \quad (B-5)$$

10. Morrison, D., Multivariate Statistical Methods, Chapter 3, McGraw Hill, N.Y.

where $D = \sigma_{xx}\sigma_{yy} - (\sigma_{xy})^2$ and c is a constant. If we approximate Δx and Δy by Gaussian random variables, then the probability that the target solution lies within the ellipse (B-5) is $1-e^{-c/2}$. To show this, we rotate the coordinate system so that Δx and Δy are replaced by variables z_1 and z_2 along the principal axes of the ellipse. Equation (B-5) becomes

$$\frac{z_1^2}{\sigma_1^2} + \frac{z_2^2}{\sigma_2^2} = c \quad (B-6)$$

and the joint probability density is

$$\frac{1}{2\pi\sigma_1\sigma_2} e^{-\frac{1}{2} \left(\frac{z_1^2}{\sigma_1^2} + \frac{z_2^2}{\sigma_2^2} \right)} \quad (B-7)$$

By changing to polar coordinates and integrating (B-7) with $r \triangleq \sqrt{\frac{z_1^2}{\sigma_1^2} + \frac{z_2^2}{\sigma_2^2}}$ from $r=0$ to $r=\sqrt{c}$, we find the probability to be $1-e^{-c/2}$. Typically, c is taken to be 1 which yields a confidence region of 61%.

COMPUTATION OF Q

Equation (8) provides an expression for the covariance matrix of the state variables arising from the measurements of a single random variable. For several such independent variables w_1, w_2, \dots, w_M , the linearization (4) becomes a sum over M terms. Since these errors are independent, the covariance matrix resulting from their simultaneous presence is

$$Q^T = \sum_{i=1}^m Q(\sigma_{w_i}^2) \quad (B-8)$$

where $Q(\sigma_{w_i}^2)$ is given by (8) with σ_w^2 replaced by $\sigma_{w_i}^2$. In particular, the variance of the state variables is (approximately) linear in the variances of the measured parameters. The standard deviation is the root mean square of the individual contributions.

A plot produced by the error estimation software appears in Figure B-1a. Although not shown in that figure, the error covariance Q^{PP} is a decreasing matrix function of the data (i.e., the corresponding error ellipses "decrease" with the addition of each data point). As previously mentioned, this is a necessary consequence of the fact that the least-squares weights were chosen inversely proportional to the true variances. On the other hand, we note that the ellipses plotted in Figure B-1a are not always "decreasing." Such behavior results from the propagation of the error along the target track; the covariance at a given time being a complex function (equation (B-3)) of the velocity covariances Q^{PV} and Q^{VV} as well as Q^{PP} . In addition, one can expect a discontinuity between the two target segments since, at the time of maneuver, the number of state variables increases by two.

Also included in the plot are points computed by an MLE tracking routine from simulated data. The weights a_t and b_t of equation (1), used in the tracker were chosen to be $(\sigma_\theta)^{-2}$ and $(\sigma_f/f_0)^{-2}$, respectively. Typically these are *a priori* variances and hence not necessarily very accurate. However, the error model is valid for any choice of these weights, independent of the measurement statistics; those used in the model should simply be the same as those in the cost function of the tracker. On the other hand, the predicted errors still depend critically on the statistics through the variances $\sigma_{w_i}^2$ of equation (B-8). In practical situations these must be estimated from the data. For the tracker used here, the variances for the bearings and frequencies were determined by averaging the errors, $(\theta_i - \tilde{\theta}_i)^2$ and $(f_i - \tilde{f}_i)^2$, with θ_i and f_i determined by the solution (not true) track. The averaging also took into account the reduced number of degrees of freedom resulting from the dependence of the θ_i and f_i forced by equation (10). The results were used to plot the error ellipses in Figure B-1b. (Those in Figure B-1a used the true track and true $\sigma_{w_i}^2$, i.e., σ_θ^2 and σ_f^2 .)

Note that the $\sigma_{w_i}^2$ estimates for the southern-most ellipse were apparently too small. This is not surprising considering the small number of data points (three) involved in the averaging. No solution was found for less than three data points since the number of state variables (five before the maneuver) outnumbered the number of equations ($2 \times$ number of data points for frequency and bearing).

As a second example, Figures B-2a and B-2b illustrate the effects of an uncertainty in platform velocity. This case is identical to Figures B-1a and B-1b except that the bearing and frequency measurement errors have been set to zero, while errors in platform velocity have been introduced. In particular, $\sigma_{w_1}^2 = \text{var}(v_x)$ (not of v'_x !), and $\sigma_{w_2}^2 = \text{var}(v_y)$ with the values chosen to be $\sigma_{w_1}^2 = \sigma_{w_2}^2 = 0.5(\text{knot})^2$ so that the RMS speed error is 1.0 knot. These errors correspond to errors in the parameters z of equation (9) rather than in c as in Figure B-1. Also, note that the least squares weights (taken the same as in Figure B-1) are not equal to the inverses of σ_θ^2 and $(\sigma_f/f_0)^2$, both of which are zero in the current sample.

	Initial time (minutes)	Final time (minutes)	Course (degrees)	Speed (knots)
Target leg 1	0	33.33	315.0	$ \tilde{v}'_1 = 15.0$
Target leg 2	33.33	72.0	10.0	$ \tilde{v}'_2 = 15.0$
Platform 1	0	72.0	30.0	6.0
Platform 2	4.0	64.0	135.0	5.0

State Variables: $P'_{ox}, P'_{oy}, \tilde{v}'_{1x}, \tilde{v}'_{1y}, \tilde{v}'_{2x}, \tilde{v}'_{2y}, f_0$

Table for Figures B-1a and B-1b.

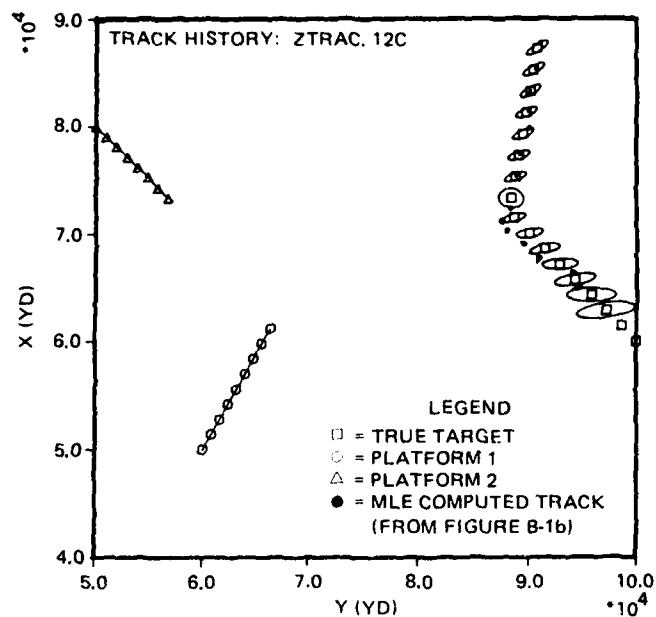


Figure B-1a. Comparison of predicted errors with actual results from an MLE tracking routine. The bearing standard deviation is $\sigma_\theta = 1.5^\circ$ and frequency standard deviation is $\sigma_f = 0.01$ with a base frequency of 100 Hz. The MLE weights are σ_θ^{-2} and $(\sigma_f/f_0)^{-2}$. The error ellipses, etc., were generated by the sensitivity error analysis routine from the true track. (See preceding table for added information.)

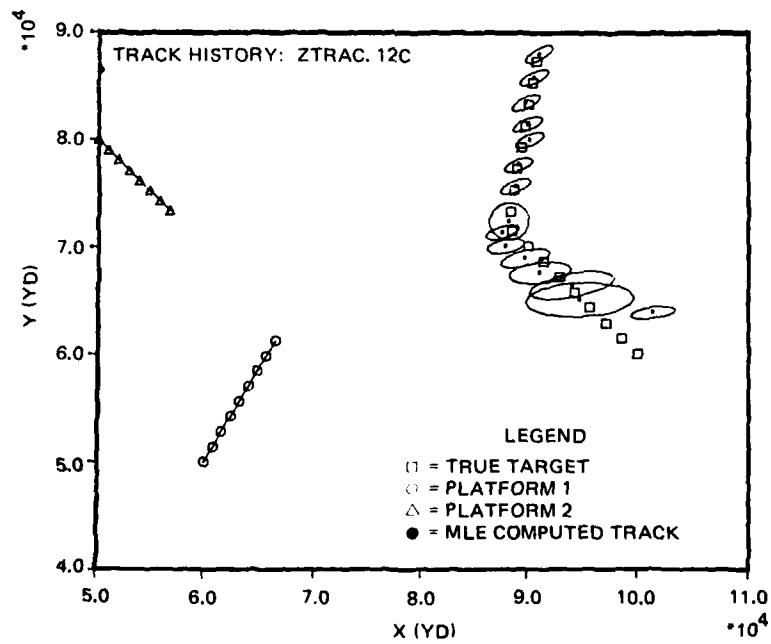


Figure B-1b. The "reverse" of Figure B-1a. This plot was generated using an actual MLE tracking routine on simulated data. Pseudo random noise with standard deviations of 1.5° and 0.01 Hz, were added to the true bearing and frequency values. The error analysis producing the ellipses used estimates of measurement errors and target positions supplied by the MLE algorithm. (See preceding table for added information.)

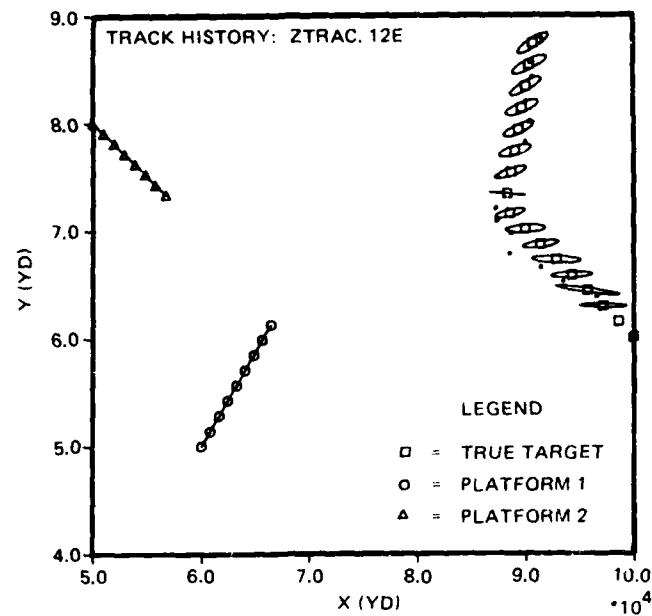


Figure B-2a. Sensitivity due to 1.0 knot error in platform velocity generated from the true track. See text for complete description.

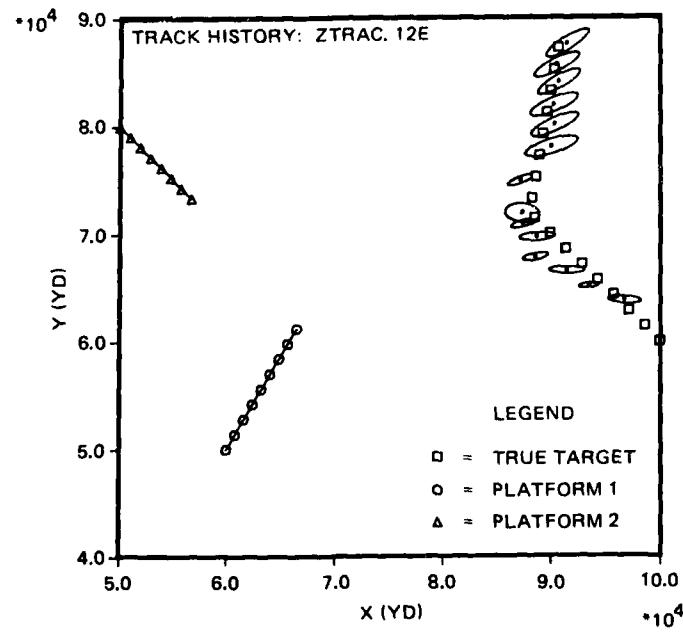


Figure B-2b. Same as B-2a but using an MLE tracking routine with simulated data.

APPENDIX C

SOME THEOREMS ON POSITIVE DEFINITE MATRICES

THEOREM I

If A and B are positive definite symmetric matrices satisfying $A \geq B$, then $B^{-1} \geq A$.

The proof relies on the following lemma:

Lemma I ([9], pg 307)

If A and B are symmetric matrices with A positive definite, then there exists an (more than one) invertible matrix H such that $H^\dagger A H$ and $H^\dagger B H$ are diagonal and real (\dagger indicates transpose).

pf of Theorem I

Given any vector x , let $y = Hx$. Then $y^\dagger A y \geq y^\dagger B y$ implies

$$x^\dagger H^\dagger A H x \geq x^\dagger H^\dagger B H x \text{ for all } x \quad . \quad (C-1)$$

Let D_A and D_B be the diagonal matrices of Lemma I. (C-1) implies

$$x^\dagger D_A^{-1} x \leq x^\dagger D_B^{-1} x \text{ all } x \quad . \quad (C-2)$$

Since the entries of D_A and D_B must be positive,

$$x^\dagger D_A^{-1} x \leq x^\dagger D_B^{-1} x$$

or

$$x^\dagger H^{-1} A^{-1} H^\dagger x \leq x^\dagger H^{-1} B^{-1} H^\dagger x \text{ for all } x \quad . \quad (C-3)$$

Given any z , let $x = H^\dagger z$, then

$$z^\dagger A^{-1} z \leq z^\dagger B^{-1} z \text{ all } z \quad . \quad (C-4)$$

which proves the theorem.

THEOREM II

Let A and B satisfy the hypotheses of Theorem I; then their largest and smallest eigenvalues satisfy $\lambda_A^{\max} \geq \lambda_B^{\max}$ and $\lambda_A^{\min} \geq \lambda_B^{\min}$.

9. Parlett, B., The Symmetric Eigenvalue Problem, Prentice Hall, Inc., N.J., p 307, 1980.

pf of Theorem II

$$\lambda_A^{\max} = \max_{\|x\|=1} x^T A x \geq \max_{\|x\|} x^T B x = \lambda_B^{\max} \quad (C-5)$$

$$\lambda_A^{\min} = \min_{\|x\|=1} x^T A x \geq \min_{\|x\|} x^T B x = \lambda_B^{\min} \quad (C-6)$$

THEOREM III

Let A and B be as in Theorem I. Then for any given $c > 0$ the concentration ellipsoid $E_A = \{x \mid x^T A x \leq c\}$ lies within the ellipsoid $E_B = \{x \mid x^T B x \leq c\}$.

pf of Theorem III

Let x lie on the surface of ellipsoid E_A . Then $c' \triangleq x^T B x \leq x^T A x = c$. Let $r = \sqrt{c/c'}$ > 1 . Then $(rx)^T B(rx) = c$. Thus rx lies on the ellipsoid E_B . Then by convexity $x \in E_B$ and hence $E_A \subset E_B$.

REFERENCES

1. Jazwinski, A., Stochastic Processes and Filtering Theory, Academic Press, N.Y., 1970.
2. Adams, W.B., "A Non-Linear Maximum *a posteriori* Estimator for Passive Ranging and Using Bearing and Frequency Measurements on a Stable Line," General Electric Co., Syracuse, N.Y., Sep 1975.
3. Powell, M., "A Survey of Numerical Methods for Unconstrained Optimization," SIAM Review, Vol 12, pp 79-97, Jan 1970.
4. Bard, Y., "Comparison of Gradient Methods for the Solution of Nonlinear Parameter Estimation Problems," SIAM J. Num. Analysis, Vol 7, pp 157-185, March 1970.
5. Van Trees, H.L., Detection, Estimation, and Modulation Theory, Part I, Wiley, New York, N.Y., 1968.
6. Taylor, J., "The Cramer-Rao Estimation Error Lower Bound Computation for Deterministic Nonlinear Systems," Proceedings of 1978 IEEE Conference on Decision and Control, pp 1178-1181.
7. Chou, S.I., "Some Drawbacks of Extended Kalman Filters in ASW Passive Angle Tracking," ONR Symposium on Advances in Acoustic Passive Tracking, NPS-62 TS 77771, Naval Post Graduate Schoo^l, Monterey, CA, May 1977.
8. Eykhoff, P., System Identification, Chapter 6, John Wiley & Sons, N.Y., 1974.
9. Parlett, B., The Symmetric Eigenvalue Problem, Prentice Hall, Inc., N.J., p 307, 1980.
10. Morrison, D., Multivariate Statistical Methods, Chapter 3, McGraw Hill, N.Y.

END
DATE
FILMED

10-81

DTIC